

Chapter 3

Dilemmas For Programs, Sponsors And Evaluators



Using Outcome Evaluation To Guide Grant-Making: Theory, Reality And Possibilities

This is an edited version of an article originally printed in *Nonprofit and Voluntary Sector Quarterly*, 29, pages 330-334, 2000.

Doug Easterling, PhD
University of North Carolina at Greensboro
Center for Study of Social Issues

Raising The Bar For Evaluation

Over the past few years, the philanthropic sector has undergone a critical redefinition of grant making. Grants are seen less and less as *gifts* or *contributions* than they are as *investments*. This shift in terminology is not just semantic; foundations now have much higher expectations as to what nonprofit organizations will do with their grants. The sector is being counseled to pay deliberate attention to the *value to society* that is produced by a grant (Porter and Kramer, 1999; Prager, 1999) and even to step into the role of an *active* investor (i.e., a “venture capitalist”) in order to push grantee organizations to their full potential (Letts, Ryan and Grossman, 1997).

The grants-as-investments paradigm immediately raises the issue of evaluation: foundations need to be able to measure what grantees accomplish with their grants in order to compute the return on their investment. Moreover, foundations operating under the new paradigm are much more impressed with *outcome* evaluation (i.e., an objective assessment of the actual effects of the funded program on the target population) than they are with *process* evaluation (i.e., an assessment of how the program was delivered by the staff and received by the clients).

This increasing emphasis on outcomes and results is reinforced by a number of other trends operating within the philanthropic and nonprofit sectors. For example, in the wake of some highly publicized cases where nonprofit executives paid themselves excessive salaries and exercised questionable spending practices, we have seen movement within Congress, foundations and the general public to require more accountability from the sector. In addition, many of the new foundations emerging over the past decade have been created out of the sale of hospitals and health systems (Milbank Memorial Fund, 1999). Such “conversion foundations” are often managed by health administrators, who are accustomed to making decisions that are based on concrete performance indicators.

Under this new results-based paradigm, funders are no longer mollified by *activity-based* progress reports that count how many meals were served, how many beds were filled, how many addicts were treated, or how many offenders showed up for group counseling.

Grantees now must gather data that accommodate bottom-line thinking, showing the concrete, tangible changes that resulted from the foundation’s support.

The Sober Reality

Although foundations are becoming increasingly clear with their expectation for evaluation, the task facing grantees is often clouded and confused (Fine, Thayer and Coghlan, 1998; Walker and Grossman, 1999). Foundations will not elicit rigorous outcome evaluations from their grantees simply by asking for them. For many nonprofits, the requirement to conduct evaluation appears extraneous to the core mission and activities of the organization. Even those organizations that do recognize the value of evaluation often find it difficult to build this new function into their existing operations. And those organizations that actually do evaluate their programs rarely obtain the type of data that would allow funders to compute the return on their investment or the ratio of benefits to costs.

The requirement to conduct outcome evaluation appears as a huge order for most nonprofit organizations (Billitteri, 1998). To support the transition from informal to rigorous evaluation, some funders have devoted resources to the cause of building evaluation capacity within the nonprofit sector. United Way of America produced the guidebook, *Measuring Program Outcomes: A Practical Approach*, which is designed to assist local United Ways and their funded organizations in setting up useful evaluation systems. Foundations such as W.K. Kellogg and Skillman have prepared their own manuals (*W.K. Kellogg Foundation Evaluation Handbook* and *A Guide to Evaluation for Skillman Foundation Grantees*), which cover topics ranging from logic models to working with research consultants. More and more workshops and conferences are devoted to training the nonprofit and government sector in evaluation methods. Foundation-supported websites such as Innonet (www.inetwork.org) provide users with

guidance and exercises that help even novices design an evaluation of their program, and then allow the user to upload those designs for on-line feedback.

All of this exposure around evaluation has raised expectations and anxiety among program managers, without necessarily increasing the quantity or quality of the evaluation work that is actually performed. If a foundation is to employ outcome evaluation as a tool in its quest for organizational effectiveness, the practical limitations of this tool must be acknowledged and addressed. The following issues have emerged as critical limitations as nonprofit organizations have ventured more deeply into outcome evaluation.

Outcome Evaluation Is Expensive

Many evaluation manuals and workshops begin with the reassuring axiom that evaluation is simply an extension of normal, adaptive behavior. For example, Posavac and Carey (1997) exhort that, “the practice of evaluating one’s own efforts is as natural as breathing” (p.1). These

statements are intended to demystify the field of evaluation and to convince program managers that they can conduct evaluation even if they don’t have formal training in statistics. Although some forms of evaluation (particularly *formative* evaluation) can be built directly into an organization’s operations with relatively little effort and cost, rigorous *outcome* evaluation rarely proves to be a natural, inexpensive or simple addition to a program. Measuring outcomes and impacts requires questionnaires, interviews, tracking clients



after they leave the program, data management and analysis: tasks that impose a real burden on staff and clients. In addition, few nonprofit organizations have the expertise on staff to design rigorous studies, identify valid instruments and carry out sophisticated statistical analyses. Workshops, training manuals and consultation can be effective in building basic evaluation competencies, but a good, solid outcome evaluation will invariably require the addition of new staff (or dedicated consultants) who bring extensive training in evaluation methods.

The Outcomes Of Social Programs Are Elusive

Few of the programs funded by foundations can be evaluated with the degree of precision that foundation board and staff would like. Particularly for board members who are accustomed to reviewing quarterly financial returns to evaluate the performance of their investment managers, the ambiguity surrounding the assessment of *programmatic* outcomes can be very disconcerting.

One problem related to elusiveness is that many foundation-funded programs have “soft” outcomes.

Indeed, nonprofit managers often contend that evaluators and their methods “miss” the positive effects of community-based programs (“I know my program works, regardless of what the evaluation shows.”) Although many “soft” outcomes (e.g., self esteem, self sufficiency, community mobilization) can be measured reliably and validly, evaluators do face challenges in trying to fully capture the effects of programs that address complex issues such as substance abuse, violence, homelessness, discrimination and social injustice.

A parallel limitation in applying outcome evaluation to social programs is that many of these programs have *lagged* outcomes (occurring well into the future). Especially within the field of prevention, even the programs that are effective have benefits that do not occur until years after the intervention has been delivered. Take the case of David Olds’ home-visitation program, which has attracted considerable attention from foundations and policy makers because of its demonstrated effectiveness in reducing the likelihood of violence among children of mothers who were visited by nurses during pregnancy and early childhood; the reductions in youth violence were observed 15 years after the intervention (Olds, et al., 1998). Significant expenses and complications are involved in following clients even a year after they exit a program.

Another methodological complication associated with outcome evaluation is that most of the programs funded by foundations operate in “noisy” systems. In other words, at the same time that a client is receiving some service from the grantee organization, that client is also experiencing a host of extraneous influences.¹ In some programs (particularly school-based programs), the clients undergo a whole host of developmental changes at the same time that they are being influenced by the intervention; any observed change in behavior or knowledge might be due to the intervention, but it might just as well be attributed to a *maturation* process. Likewise, *secular trends* (e.g., a changing economy, new technologies, a deteriorating neighborhood) also exert an influence on the clients of a program, over and above the impact of the intervention.

Because of these “background effects,” the effectiveness of a program cannot be evaluated simply by assessing changes in client behavior or well-being; a true outcome evaluation requires that we identify the *unique contribution* of the program on the client. As every evaluation textbook makes excruciatingly clear, the challenge for program evaluators is to estimate what would have happened to the program’s clients in the absence of the intervention. In theory, approaches such as random assignment, a “matched” comparison group and interrupted time-series designs offer solutions to this dilemma, but these strategies are exceedingly difficult to implement within most nonprofit settings, particularly given the level of resources typically available for evaluation.

Grantees Are Resistant To Outcome Evaluation

So far, we have been focusing on the *technical* limitations of relying on outcome evaluation as a tool for measuring the effectiveness of funded organizations. The *political, sociological* and *psychological* issues are just as critical. In particular, any attempt to introduce outcome evaluation into nonprofit organizations must contend with the pervasive sense of threat (and even dread) that accompanies the word “evaluation.” Despite all its potential as a tool for learning, evaluation is traditionally viewed solely as a yardstick for measuring performance, with all the associated connotations of report cards, judgmentalness and punishment. Particularly when the evaluation is driven by the funder, and the funder emphasizes outcomes, the common response among grantees is fear and

avoidance.

Increasing The Relevance And Utility Of Evaluation Among Grantees

Even though outcome evaluation is expensive, complicated, threatening and inherently political, this does not necessarily mean that it is inappropriate; these qualities are true of most tools that facilitate change and growth. Evaluation has the power to bring a more deliberate, informed mode of decision making to foundations and the organizations they fund. However, for evaluation to achieve that potential, those who fund and practice evaluation need to be more strategic, thoughtful and sensitive in their approach to evaluation.

The following recommendations are intended to advance the state of evaluation within the philanthropic sector.

Set Clear, Reasonable Expectations For Measuring Outcomes

A good deal of the frustration associated with evaluation can be overcome through a more intentional definition of what the program is intending to accomplish, which in turn allows the measurement to be more focused. In particular, when foundation staff and boards are more clear and specific as to the intent of a grant or an initiative, the evaluation task becomes more straightforward. Evaluation methods can capture program-induced changes, but only if those desired changes are articulated beforehand.

On the other hand, even with a clear set of objectives, the evidence gathered by an evaluation may not be as definitive as the board and staff are hoping. A precise evaluation of outcomes requires significant funding, long-term follow-up, and random assignment of clients and/or grant



applicants: provisions that run against the grain of traditional foundation practice. It is important to recognize and appreciate the various costs involved in moving from an evaluation that yields *suggestive* evidence to an evaluation that yields *definitive* evidence. Once a decision is made as to the scope of the evaluation design, the foundation should accordingly adjust its expectations for outcome data.

Use Logic Models

The outcomes of strongest interest to foundations and funded organizations tend to be the “hard” effects that occur down the road: graduation from high school, staying in a job long run, maintaining healthy behaviors, improved social fabric throughout a community. The typical one-shot evaluation is inherently unable to detect these long-term outcomes. This issue of latency can be addressed to some extent by measuring shorter-term outcomes (e.g., increases in knowledge, new skills) that serve as proxies for the ultimate outcome of a social program (e.g., avoiding unhealthy behavior such as smoking or substance abuse).

Logic models connect the shorter term and longer term outcomes together in a chain of causality (Weiss, 1995). Logic models are very useful in establishing a common set of expectations as to how a program will unfold, if it is successful. By laying out the logic model at the beginning of the program, those stakeholders who are predisposed to assessing success in terms of a “bottom line” have a conceptual tool that allows a more reasonable scorecard for tracking the program’s progress.

Involve Grantees As Partners In The Specification Of Desired Outcomes

Although foundations clearly have a right to define the expectations and objectives of their grants, the organizations receiving funding are not passive actors in this exchange. Indeed, nonprofits arguably have an even larger stake in establishing the criteria by which they will be held accountable by their funders and by themselves. These two perspectives must be brought together through open-ended funder-grantee negotiations, which in turn requires that both players share their own bottom lines, listen carefully to the other’s perspective and then seek common ground. The process is complicated by the power differential that invariably enters into any conversation between funder and grantee, particularly when defining the terms and expectations of a grant. The degree to which the foundation and grantee succeed in establishing mutually agreed-upon outcomes depends largely on the strength and integrity of their relationship, which in turn rests on their mutual trust, respect and appreciation for one another’s role.

Build Evaluation Capacity At The Same Time The Evaluation Is Carried Out

Foundations typically enter into the evaluation arena because of their desire to know the effects of their grants. At the same time that the foundation is meeting its own need to learn of the effects of its grants, it can serve its grantees by building their capacity to conduct evaluation.

Building the evaluation capacity of nonprofit organizations requires a concerted, long-term strategy on the part of funders, evaluators and the organizations themselves. One-shot workshops and introductory manuals can succeed in sensitizing an organization to the possibilities associated

with evaluating its activities, but these resources do not, by themselves, infuse deeper levels of skills and knowledge, build data systems, or transform organizational culture in a way that allows evaluation to take root. Thus, evaluation rarely succeeds in attaining its ultimate goal of improving the design and implementation of the program being evaluated.

Although much has been written about building evaluation capacity among community-based organizations (e.g., Fetterman, Kaftarian and Wandersman, 1996), there are two critical factors that come into play when a foundation works with a grantee. The first is fear. Nonprofit managers tend to be skeptical and suspicious of evaluation, particularly evaluation that is imposed by an external agent. To undo those connotations, a deliberate effort must be made to bring out intrinsically meaningful evaluation questions among the staff and board of grantee organizations. The intent is to transform evaluation into a tool for self-directed learning for the grantee.

A second major factor that comes into play in building the evaluation capacity of grantees is the potential disconnect between evaluation-related activities (e.g., data collection, analysis, interpretation, learning) and program-related activities (recruiting staff and clients, developing work plans, writing grants, gaining public exposure). To be successful in introducing evaluation into a nonprofit organization, the evaluators and data systems must accommodate the existing culture and constraints, at least at the beginning. At a bare minimum, this means that the foundation must make available resources that are sufficient to bring in the necessary staff, equipment and software. Further, the staff of the organization needs to have a strong say in how the system for evaluation will fit within the system for

delivering services. This requires time, flexibility and sensitivity to different perspectives. Most of all, the staff must be assured that their own interests are being met in pursuing evaluation.

Increasing the Power of Evaluation within Philanthropy

Focusing evaluation on the outcomes of grantees leaves the foundation without a sense of how its own decisions are determining those outcomes. The return that a foundation achieves on its grants obviously depends on the success of its grantees. It is important to recognize how much that success depends on the foundation's approach to grant making, including the foundation's choice of funding areas, the duration and size of grants, the foundation's style of interacting with grantees, the requirements imposed on grantees, and the availability of technical support. Unless the foundation evaluates the effect that those decisions and activities have on the performance of grantees, no amount of grantee-level outcome evaluation will move the foundation to the point of realizing its own potential.

The paucity of foundation-focused evaluation severely limits the effectiveness of the philanthropic sector. No foundation can achieve its true potential without the deliberate data-collection, analysis and reflection that are the keys to good evaluation. Without objective data describing the results of its actions, a foundation is unable to learn from its mistakes and its successes. Foundations expect grantees to push themselves to increase relevancy, effectiveness, and wisdom; grantees should expect just as much from funders.

References

- Billitteri, T.J. (1998). "Tracking the effectiveness of charities." The Chronicle of Philanthropy, 10(16), 35-36.
- Easterling, D. and Csuti, N.B. (1999). "Using evaluation to improve grant making." Grantmaker's Evaluation Network 6, (Summer), 2-3.
- Fetterman, D.M., Kaftarian, S.J., and Wandersman, A. (1996). Empowerment evaluation: Knowledge and tools for self-assessment and accountability. Thousand Oaks, CA: SAGE Publishers.
- Fine, A.H., Thayer, C.E. and Coghlan, A. (1998). Program evaluation practice in the nonprofit sector. Washington, DC: Innovation Network, Inc.
- W. K. Kellogg Foundation (1998). W.K. Kellogg Foundation Evaluation Handbook. Battle Creek, MI: W.K. Kellogg Foundation.
- Fine, A.H., Thayer, C.E. and Coghlan, A. (1998). Program evaluation practice in the nonprofit sector. Washington, DC: Innovation Network, Inc.
- Letts, C.W., Ryan, W., and Grossman, A. (1997). "Virtuous capital: What foundations can learn from venture capitalists." Harvard Business Review, 75 (March/April), 36-44.
- McNelis, R.H. and Bickel, W.E. (1996). "Building formal knowledge bases: Understanding evaluation use in the foundation community." Evaluation Practice, 17(1), 19-41.
- Milbank Memorial Fund (1999). New foundations in health: Six stories. New York: Milbank Memorial Fund.
- Olds, D., Henderson, C.R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettit, L., Sidora, K., Morris, P., and Powers, J. (1998). "Long-term effects of nurse home visitation on children's criminal and anti-social behavior: 15-year follow-up of a randomized controlled trial." Journal of the American Medical Association, 280, 1238-1244.
- Patrizi, P. and McMullan, B. (1998). "Realizing the potential of program evaluation." Foundation News and Commentary, 40 (3), 30-35.
- Porter, M.E. and Kramer, M.R. (1999). "Philanthropy's new agenda: Creating value." Harvard Business Review, 77 (November/December), 121-130.
- Posavac, E.J. and Carey, R.G. (1997). Program evaluation: Methods and case studies (5th edition). Upper Saddle River, NJ: Prentice-Hall.

Prager, D.J. (1999). Raising the value of philanthropy: A synthesis of informal interviews with foundation executives and observers of philanthropy. Washington, DC: Grant makers in Health.

United Way of America (1996). Measuring program outcomes: A practical approach. Alexandria, VA: United Way of America.

Walker, G. and Grossman, J.B. (1999). Philanthropy and outcomes: Dilemmas in the quest for accountability (P/PV Briefs). Philadelphia, PA: Public/Private Ventures.

Weiss, C. H. (1995). "Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families." In Connell, J.P., Kubisch, A.C., Schorr, L.B., and Weiss, C. (eds.), New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts. (Washington, DC: The Aspen Institute).

Evaluation Dilemmas For Grantees

Christine Patterson, MSW, ACSW, LCSW
Arkansas Department of Health
Office of Minority Health

Dilemmas regarding evaluation for community-based programs are numerous, with the core driving force being that evaluation must be done. Grassroots leaders and residents characterize funders as being unresponsive and non-inclusive. Communities are concerned with ownership: Who will control the data once it is collected? What resources will develop as a result of the evaluation? During my years of experience, the one consistent element that has been evidenced as a requirement to being successful in working with communities of color is relationship building. These communities must have a sense of commitment from funders before they are willing to be a collaborative partner.

First, community groups generally have not done evaluations; they have been evaluated. The unfortunate history has been that they have been evaluated with rigor only when someone in power decides that there is a "problem".

Most community-based groups do not feel that they have "permission" to learn from their errors or to shape or improve a program through iteration. They have to get it right, and get it right the first time or lose their money. In this sense, the normal request for proposal phrase, "Funding for subsequent years in the grant period is dependent on demonstration of adequate performance in the current year," is delightfully clear. We have the money and will unilaterally make the life and death decisions for your program.

Even when it is not a "gotcha game," evaluation too often remains almost entirely externally driven, with little internal value or utility to the community group. The basic process of evaluative assessment is useful in looking at some of these issues. First, doing a meaningful evaluation requires an agreement on objectives and expectations for the program and clarity in

the internal logic of the program's model. Talk to the stakeholders in most community-based initiatives and you almost always discover a range of expectations that is awe-inspiring. When the program involves collaboration among disparate organizations, agencies, and professionals in a community, ambiguity may be the enemy of continued agreement and collaboration.

Second, community organizations tend to see grant funds as core operating money, not as special project funding. Many community-based organizations do not have a regular appropriation for core program activities, but instead have to juggle a hodgepodge of grants and other funding so they can keep their doors open, staff paid, and their services going. Raising the issue of evaluation in this context is not a welcoming subject and in fact can be perceived as a threat to the continued operation and survival of the organization.

There are all kinds of real methodological issues that most community-based organizations don't have the technical evaluation and statistical skills to cope with.

If funders really want grassroots groups to be able to submit serious and substantive proposals, then they must give 90 or 120-day notice periods. A 30-day request for proposals that seeks a complex collaboration that brings together grassroots groups, statewide health stakeholders, academic medical centers, and local health professions is not being realistic. Couple this short time period with an expectation for an effective evaluation model to be in place, and it is no wonder that many community-based organizations do not attempt to submit applications.

All of my colleagues supported the inclusion of community-based organizations in the initial phase of any project from planning to implementation to evaluation. This model encourages local participation and buy-in into the research process and facilitates the sharing of local knowledge. Therefore, evaluation will be seen less as a threat and more as a means of ensuring that programs and services are working as planned to effectively meet the needs of the target population. By reducing the occurrence of these dilemmas, the relationships between funding agencies and grantees can develop in healthier ways that benefit both sides.

A Community Dilemma: Community Autonomy Versus The Need For External Funding

*Christine Patterson, MSW, ACSW, LCSW
Arkansas Department of Health
Office of Minority Health*

A health promotion effort was needed in a community located in Newport News, Virginia. Approximately 98% of the community residents were African American, with limited access to health care and health related services. The highest incidence and prevalence for diseases and illnesses for the city were concentrated in this community. The socioeconomic status of the

community was the lowest for the city, with the highest rates of children and families living below the poverty level.

Several public and private health and human services organizations formed an alliance to promote the health and well being of the children and youth in the community. The Alliance started to plan services for the community without approaching or involving the citizens who resided in the community. The community felt left out of the process, and expressed the opinion that they should determine what programs; services and resources should be brought into their community. The citizens of the community became upset and insisted on being involved in planning the services.

The major dilemma was this: the citizens did not have the resources to implement programs and services without the support of the Alliance. How could the citizens get support from the Alliance, without alienating the Alliance members?

The resolution was to plan collaboratively to establish a multidisciplinary center providing services to the residents. After three years of planning, in a process that included the city government, the Alliance and community residents, a multidisciplinary youth center was relocated from a school site to an old, renovated recreational center. The youth center formed its own board of directors, comprised mainly of citizens and administrators from schools in the community. The board created a slot for the principal of the school where the program was originally located. When that principal was reassigned, a heated discussion centered around whether the position was created for an individual or a representative from the school. The organizations and agencies continued to function as an Alliance and provided resources and services to the youth at the center. The board of the youth center divorced itself from the Alliance after about 5 years and developed contracts with individual Alliance organizations and agencies for programs and services. Today the center continues to provide a variety of educational, recreational and prevention programs. The Alliance continues to function with prevention programs, "Youth Development-Developmental Assets" for the entire city. Recently the city formed an Office of Youth at the urging of the Alliance.

An Evaluator's Dilemma: How Can I Collect Comparable Evaluation Data Across Program Sites, When Some Sites Are Truly Unique?

*Ross F. Conner, PhD
University of California at Irvine
School of Social Ecology*

I have to admit, I finally met my match. As part of a larger evaluation of nine Healthy Community projects, I arrived in South El Paso expecting to lead a focus group meeting. When I walked through the door, I discovered:

- Thirteen 8-year-old boys on a soccer team, in uniforms and with soccer balls bouncing;
- Twelve teenage high school boys who were members of a break-dancing team;
- Two teenage girls who had just arrived on the scene, wearing their cheerleading uniforms;
- A set of adults involved in one community project;
- Another set of adults involved in a different project; and
- About a dozen staff members.

The kids spoke English; the adults spoke Spanish. The scene was utter confusion. It was 5:30pm and prime dinnertime for the 8 year-olds, who had just been outside playing soccer--and there was no food anywhere.

I had come to gather data for an evaluation, expecting to use focus group discussions, a method well suited to the evaluation questions, the project and the community. Neither my experiences nor any of my training had prepared me to deal with such a large, diverse group--and diverse along so many dimensions. The adults were ready for a respectful discussion with this "doctor from California;" the staff were ready to talk about their programs; the kids were ready for dinner.

Well, first things first. We ordered food--quickly and lots of it: 15 pizzas and cokes. Then, I split the crowd into two groups, with the adults in one group and the kids, along with the majority of the staff, in another. A foundation officer along for the visit took the adult group and conducted a discussion in passable Spanish. I took the kids to a different room and tried to lead a discussion in English. Given the reactions I got, I might just as well have been speaking Chinese. It was pretty clear that the main focus of the teenage boys was those two cheerleaders. The teenage boys, when able to get their eyes away from the cheerleaders, acted like teenage boys everywhere: I was lucky to get one-word answers to any of my questions. Nothing even close to a discussion ever occurred. The two teenage girls were delighted to talk at length -- about anything. The 8-year olds...well, let's say that the phrases "have a discussion" and "8 year-olds" don't fit together at all.

These little guys were unbelievably well-behaved and polite, and they at least tried to respond. They talked more than the teenage boys, who, true to form, completely ignored the presence of these little guys.

I prayed for the pizza man to arrive. He finally did--and brought an instant, spontaneous end to the "discussion." The pizzas disappeared in a flash. The evening ended on a very positive note: the staff members stayed around after dinner, and we had an rich discussion, out of which came lots of data and many insights.

Dilemmas For Programs, Sponsors And Evaluators

*Doug Easterling, PhD
University of North Carolina at Greensboro
Center for Study of Social Issues*

The Macro-Micro Dilemma.

Foundations and government agencies often fund multiple sites under a common initiative (e.g., REACH 2010). While acknowledging that the sites are all different, the funder still would like to evaluate the initiative using a common metric. This impetus typically leads to one of two distinct strategies, each of which creates a dilemma for the funded agencies:

1. *Establish a set of “common” outcomes to be measured across all sites.* For example, all programs funded under a hypertension-prevention initiative might need to measure their clients in terms of changes in diet, exercise and monitoring of blood pressure. Imposing these core measures on all sites will restrict the creativity of the funded projects, weeding out ideas that do not comport with the funders’ implicit definition of what needs to be included in an “effective” project. This conformity in programming may have particularly detrimental consequences for programs that are attempting to improve health with individuals whose beliefs, attitudes and culture are at odds with “standard” models of behavior change. In other words, the required evaluation could force out the very strategies that might lead to the greatest improvements in health among a culturally diverse group, by building off deeply rooted (but “nonstandard”) beliefs about health, relationships, balance and the meaning of life.
2. *Assess each site according to improvements in “capacity.”* If the funder abstains from imposing a set of common evaluation measures (in order to leave open the question of what constitutes an “effective” program in any given community), the question arises as to how to evaluate the effectiveness of the initiative. Assuming that the funder wants to employ a common metric (as opposed to simply conducting distinct evaluations within each site), it is necessary to identify some outcome that should occur in common within each of the sites. Such an outcome is unlikely to pertain to the specific programs that the various sites implement, because those programs will have varying strategies, target populations and desired outcomes.

Rather, the common outcomes will likely relate to factors such as collaboration,” “community capacity,” or “program quality,” factors that have relevance within all the funded sites, given the funder’s strategy and philosophy.

This “light-handed” approach to evaluating community-driven initiatives avoids the problem of forcing programs into molds that may not fit the local culture, but it raises two new dilemmas. First, outcomes such as “community capacity” tend to be fuzzier, and thus harder to interpret than program-specific outcomes such as reductions in blood pressure or increases in the number of persons screened for diabetes. Especially with publicly-funded programs, an improvement in community capacity may not be tangible or compelling enough to elicit ongoing appropriations. The second dilemma that arises in pursuing this approach to evaluation is that the research question being answered is not necessarily interesting to the funded organizations. Learning that the initiative raised the capacity of the funded communities by an average of x points (assuming that capacity can be measured quantitatively) does little to support the needs of those funded communities. The people who develop and operate the programs are much more interested in how well their programs are achieving their own distinct objectives and what might be done to improve those programs. As a result they are best served by a different type of evaluation, one that addresses questions related to program effectiveness, not cross-site capacity building.

Ethical Dilemmas Of Evaluation

*Johnnie Bell Bunch, BS, MS, RN
Hampton University
School of Nursing*

*Paula Lantz, PhD
University of Michigan
School of Public Health*

Program evaluation and research involving human subjects raise complex ethical, legal, and social issues. For evaluators, there is a potential role conflict between their role as researcher and their role as consumer advocate. Evaluators should promote community control and participation as partners in the development, implementation and evaluation of programs to reduce racial and ethnic disparities in health. Program evaluators must have integrity and approach the evaluation task in an ethical manner.

Three principles should serve as guidelines for considering and making decisions regarding central ethical issues in evaluation research. These ethical principles include respect for all persons, beneficence, and justice.

The principle of respect for all individuals as persons is broader than simply assuring the autonomy and self-determination of the individual. The duty to maintain confidentiality and obtain informed consent from participants in the research is mandated by respect for persons (1). Respect for persons recognizes that individuals are members of a community and that their decisions may affect others in the community. It is important for evaluators to respect the person as an individual and as a community member, especially when evaluations involve value judgment about right and wrong actions, or good and bad behavior.

Informed consent to participate in evaluation research must be obtained in an honest, open and ethical manner. People who choose not to participate must be protected from negative consequences for their choice.

The principle of beneficence involves the intent to do good, to prevent injury or harm, and to promote good. Community-based professionals and participants in evaluation situations need to make thoughtful decisions to balance potential risks and benefits. These decisions may range from determining levels of interventions in a given situation or the numbers of consumers to be served in a community with limited resources.

Program evaluation decisions also require consideration of a principle of justice. Justice requires treating people fairly. The concept of “social justice” is based on doing good for the group (2,3). Systems of social justice may promote the equality of outcomes, not necessarily equality in process.

Tension may exist between people who most value social justice and people who most value autonomy for the individual, since attaining equal outcomes of goods and services necessitates forced distribution of goods and services (2). Justice requires that vulnerable people and communities not be exploited and that eligible people who may benefit from participation not be excluded without good cause. These issues become even more difficult in an evaluation context, especially when current scientific standards promote assessing the impact of an intervention by comparing a group who receives the intervention with a group who does not.

These conflicts should be resolved in partnership with individuals from the communities being served and the programs being evaluated. Racial, ethnic, and cultural differences in values may complicate situations in a multicultural mix of researchers and communities. Evaluators have a responsibility to inform themselves about the ethical, legal and policy standards that govern activities related to eliminating health disparities.

Notes

- (1) Meijer, C., Dickens, B., and Meslin, E.M. (1997). Bioethics for clinicians: Research ethics. CMAJ. 15 Volume 156(8), pp. 1153-1157.
- (2) Hall, J.K. (1996). Nursing ethics and law. Philadelphia: W.B. Saunders.
- (3) Trochim, William M. (1999). The research methods knowledge base, 2nd Edition. Internet: <http://trochin.human.cofnell.edu/kb/index.htm> (version current as of April 20, 1999).

Resolving The Dilemmas: Expectations For Evaluation

*Johnnie Bell Bunch, BS, MS, RN
Hampton University
School of Nursing*

*Paula Lantz, PhD
University of Michigan
School of Public Health*

Expectations for evaluation need to be realistic. It is important to identify potential ethical dilemmas and to plan to resolve them in partnership with the sponsors, programs and communities. Here are some considerations for setting expectations for evaluation and resolving the dilemmas.

1. Provide Useful Answers To Questions Related To Program Success

For programs designed to eliminate or reduce racial and ethnic disparities in health, an appropriate and thorough evaluation uses credible evidence to test whether the disparities are actually being reduced. Because the elimination of disparities is such a long-term process, it is important that the evaluation of these programs build in shorter-term mileposts that track progress. These mileposts show the small steps that together result in the ultimate health-status outcome. The evaluation of a community-based program should produce clear benefits:

- Building community capacity and engaging communities;
- Improving programs;
- Facilitating the planning of new or improved prevention strategies;
- Influencing policy makers and funding sources;
- Sharing what works and what does not;
- Ensuring funding and sustainability; and
- Strengthening accountability.

2. Involve Multiple Community Sectors

Community-driven evaluation means that community members involved in or affected by the program are involved. As active partners with equal voices throughout the evaluation process.

3. Listening And Learning Throughout

Evaluation provides a process by which partners involved can listen and learn. An appropriate evaluation is one that is conducted by evaluators who are culturally competent with respect to the organization that is implementing the program and the clients being served. Moreover, the design of the evaluation should be informed by the various stakeholders of the program, and thus based in the reality of the local community. Funders also have a responsibility for cultural competence in their own expectations for evaluation. Expertise regarding appropriate evaluation should flow both ways, between the sponsor and the funded organizations.

4. Define Success And Specify The Evaluation Questions

An evaluation will be appropriate and relevant only if it measures progress toward the program designers' own definition of success, and only if it incorporates the assumptions and theories of

the program designers. Programs that are truly community-based will each have their own definition of success, depending on the identified health needs, the strategies chosen, the values of the program designers, and the intended results of the initiative.

The evaluation of a program needs to reflect the theory of change that the organization is employing to bring about the desired improvement in health status. The program should articulate the theory of change. This process will help community-based organizations assure that there is an adequate linkage between their activities and their proposed outcomes. Programs will have different definitions of success, none of which may correspond exactly to the definition promoted by the funding agency. Outside evaluators must be careful not to impose their own values as to what is necessary for the health of a community to be improved.

Evaluators may have to incorporate and measure ideas beyond what is required by the sponsor, to ensure that the community's questions are answered. Also, evaluators should be open to evaluating topics that are outside their realms of expertise. If the issue is important to the community, it becomes important for the evaluation.

5. Reasonable Expectations For Outcomes

Each program should have reasonable expectations for the results that are actually possible, given the scope of the program. The cultural context of the community should be considered in defining the intended outcomes.

6. Realistic Expectations For The Program's Effect In Reducing Racial And Ethnic disparities

In order to gain real understanding of whether or not a program has been successful in promoting the health of racial and ethnic minorities, the evaluation would need to model the effects of external factors such as racism and discrimination on health.

7. Understanding The Big Picture

The evaluation of a program needs to take explicit account of the contextual factors that influence the health of the program's target population. Because the context is dynamic in nature, changes in it may influence program outcomes. Thus, it is fundamental that those involved in evaluation have a thorough understanding of the macro-level factors such as welfare reform or universal health care, which may benefit or hinder the desired successes of the interventions. At the same time, evaluations must be sufficiently grounded within the local context to provide historically and culturally sensitive interpretations of whatever changes occur under the program. It may be more efficient or effective for a program to work on changing the contextual factors that limit the opportunities of a racial or ethnic group than it is to strive for traditional notions of health behavior.

8. Politics Of Evaluation

How can the politics of evaluation be dealt with? Stakeholders are partners central to the ongoing evaluation process, and their interactions are essential to the evaluation. This process is prone to multiple and --often political- challenges: 1) establishing a common frame of reference and a cross-disciplinary vocabulary for communication; 2) identifying individual participants' agendas, both organizationally and professionally, to meet program requirements and ensure group cohesion; 3) ensuring continuity of member participation. Stakeholder interactions are dynamic

and shaped by interrelated factors, including the larger societal context, geographic area, institutional environment, power dynamics, communication styles, and decision-making styles.

9. Mutual Accountability

Accountability is an essential component in program implementation and evaluation. Because stakeholders closest to the program must report to a particular institution about the evaluation results of the program intervention, it is important to establish mutual accountability mechanisms.

CURRENT PUBLIC HEALTH PARADIGM

Health choices are not governed solely by the individual, but also by environmental and community factors over which individuals have little control.

This is the rationale for community-based strategies for disease prevention and control.